Research on Teaching Historical Data Mining and Resource Matching Mechanism of Private Data Steward

Jiyou Dong[1], Haidong Qiu[2], Jie Qiu[2,*]
[1]Yulin Normal University, Youth League Committee, Yulin, Guangxi, China
[2]Yulin Normal University, School of Artificial Intelligence, Yulin, Guangxi, China
**Corresponding Author*

Abstract: Against the backdrop of educational informatization, the traditional "one-size-fits-all" supply model of teaching resources suffers from problems such as supply-demand mismatch and uneven resource quality. Focusing on the private data steward as the core, this paper explores the mechanism of realizing accurate resource supply by combining it with teaching historical data mining. Employing literature research, case analysis, and empirical research methods, the study first analyzes the theoretical foundation of accurate supply and the current situation at home and abroad. Then, it elaborates on the functions of the private data steward, including data storage, classification management, and security assurance, as well as its applications in lesson preparation, teaching delivery, and after-class tutoring. Additionally, the paper investigates the technologies and value of teaching historical data mining, and further constructs a resource matching mechanism. Finally, it points out the challenges in terms of technology and educational concepts and proposes corresponding countermeasures, providing support for the reform of accurate supply of educational resources.

Keywords: Private Data Steward; Teaching Historical Data Mining; Resource Matching Mechanism; Accurate Supply; Educational Informatization

1. Introduction

1.1 Research Background and Motivation

In the context of educational informatization, there are significant issues in the supply of teaching resources for teachers. The traditional "one-size-fits-all" supply model ignores teachers' teaching styles, subject characteristics, and students' personalized needs, leading to supply-demand mismatch [1]. For instance, when Chinese teachers prepare lessons on ancient poetry, they need resources such as historical background materials and audio recordings by famous scholars. However, existing platforms often only provide basic text explanations, which are insufficient to meet the needs of in-depth teaching. Meanwhile, the quality of online resources varies greatly; some content contains errors or is poorly produced. These resources not only fail to support teaching but may also mislead students.

The private data steward offers a new solution to this problem. It can integrate the teaching materials accumulated by teachers over a long period, which are highly consistent with the teachers' teaching styles and the actual situation of students. Combined with teaching historical data mining technology, it records behavioral decision-making data such as teachers' resource usage frequency and students' feedback. Through algorithmic analysis, it identifies teachers' resource needs in specific teaching scenarios, enabling accurate resource recommendation, thus improving teaching efficiency and quality.

1.2 Research Value and Practical Significance

Theoretically, this study enriches the theoretical system of accurate supply of educational resources. Previous studies have mostly focused on macro resource allocation and general recommendation systems. This research, based on the individual teacher level, explores the integration mechanism of private data stewards and data mining, filling the research gap in the micro field and expanding the theoretical scope.

In practice, firstly, it improves teaching efficiency. Accurate resource supply reduces the time teachers spend on screening resources, allowing them to focus on teaching design and method improvement. Secondly, it promotes teachers' professional development. Personalized resource support helps teachers optimize their teaching strategies; through data analysis, they can identify the strengths and weaknesses of their teaching and make targeted improvements. Thirdly, it advances educational equity. The accurate mechanism enables teachers in different regions and schools to obtain suitable resources, narrowing the gap in teaching quality caused by resource disparities and benefiting more students.

2. Research Value and Practical Significance

2.1 Theoretical Foundation of Accurate Supply

The theory of educational resource allocation and the theory of personalized learning are the core theoretical supports for accurate supply. The theory of educational resource allocation emphasizes the rational allocation of limited resources to maximize benefits. In resource supply, it guides the adherence to the principle of fairness to ensure that teachers in different regions have access to basic resources. At the same time, in accordance with the principle of efficiency, resources are allocated to teachers in need through data mining, thereby improving utilization efficiency [2].

The theory of personalized learning, centered on students, requires resources to meet teachers' personalized teaching needs. By combining with data mining, the private data steward analyzes teachers' teaching strategies and resource usage habits for different student groups. For example, it pushes extended materials to teachers teaching top-performing students and basic explanatory materials to those teaching students with weak foundations, facilitating the implementation of personalized teaching.

Accurate demand analysis is the primary element of the reform. It is necessary to comprehensively consider teachers' subjects, teaching stages, teaching styles, and students' characteristics. Teachers of different subjects have significantly different resource needs: Chinese teachers need literary materials, while mathematics teachers need problem explanation materials. Resource needs also vary with teaching stages: new lessons require basic concept resources, while review lessons need comprehensive materials. Teaching styles also influence resource preferences: teachers who adopt the case teaching method need a large number of cases, and those who use group discussion teaching need guidance materials. With the help of big data and mining algorithms, teachers' potential needs can be deeply explored.

The supply of high-quality resources is the core element. High-quality resources must be accurate and scientific in content, diverse in form (documents, audio, video, etc.), and updated in a timely manner. For example, resources in the science field need to include the latest research results to meet the needs of different teaching scenarios and subject development, ensuring teaching quality.

An efficient matching mechanism is the key link. It is necessary to establish a sound resource classification and labeling system, marking resources according to multiple dimensions. Combined with algorithms such as collaborative filtering and deep learning, the matching degree between teachers' needs and resource labels is calculated to achieve accurate recommendation, thereby improving the efficiency and quality of supply.

2.2 Current Situation of Reform at Home and Abroad

Developed countries abroad have achieved remarkable results in the reform of accurate supply of teachers' teaching resources. Educational technology companies in the United States have developed platforms that collect data on teachers' usage of lesson preparation and teaching resources, as well as students' academic performance and learning behavior data. With complex algorithms, these platforms accurately push resources such as courseware, videos, and exercises to teachers, improving the efficiency of resource acquisition and adaptability. Some schools in the United Kingdom have adopted intelligent teaching assistants that monitor the teaching process in real time, analyze teacher-student data, and automatically recommend auxiliary resources such as animations and cases when students encounter difficulties in understanding.

In China, with the advancement of educational informatization, more attention has been paid to the accurate supply of resources. Some universities have built resource management platforms to integrate on-campus resources and analyze teachers' data, pushing academic papers and teaching cases. Primary and secondary schools have collaborated with enterprises to introduce recommendation systems, which push preview materials, courseware, and after-class assignments according to teaching plans, students' learning situations, and requirements of subjects and grades.

However, there are gaps between China and foreign countries: foreign countries lead in technology application and algorithms, with mature data collection and analysis systems, and can also provide a wealth of international resources. In contrast, China faces problems such as poor data quality (incompleteness, inaccuracy), the need for algorithm optimization, and weak integration and supply of international resources [3]. China can learn from foreign experience and strengthen its own construction to narrow the gap.

3. Private Data Steward: Functional Characteristics and Application Scenarios

3.1 Core Functions of the Private Data Steward

The data storage function provides large-capacity and high-reliability services, supporting the storage of files in multiple formats. Moreover, the storage capacity can be expanded according to teachers' needs, meeting the demand for storing teaching materials accumulated by teachers over the long term and preventing data loss due to insufficient space.

The classification management function combines AI algorithms and manual operations to automatically classify materials based on content, subject, etc. For example, Chinese materials can be subdivided into categories such as ancient poetry and writing guidance. Teachers can also customize classifications according to the teaching progress, greatly improving the efficiency of material search.

The security assurance function adopts multiple encryption technologies to protect the security of data transmission and storage. Only authorized teachers can access the data. At the same time, regular data backup is conducted, and teachers' privacy is protected in compliance with laws and regulations, preventing data loss due to system failures.

3.2 Diverse Applications in Teaching

In the lesson preparation stage, teachers can search for excellent historical lesson plans and courseware in the steward and optimize their designs based on the current teaching situation. The steward also recommends high-quality resources according to teachers' resource usage preferences, broadening their thinking in lesson preparation and enriching teaching content.

In the teaching delivery stage, teachers can access the steward's materials at any time through multiple terminals, quickly calling resources such as audio, video, and animations to enrich teaching forms. Meanwhile, interactive functions are used to push exercises and conduct quizzes, enabling real-time adjustment of teaching progress [4].

In the after-class tutoring stage, based on students' homework and feedback, teachers select targeted tutoring materials from the steward. For example, they push basic explanatory materials to students with weak mathematics foundations and competition questions to top-performing students. Additionally, teachers communicate with students online through the steward to achieve personalized tutoring.

3.3 Application Cases and Effects

In practice, some teachers faced problems such as disorganized data storage, low lesson preparation efficiency, and poor teaching effects before using the private data steward. After introducing the steward, they classified materials according to textbook chapters and teaching types. When preparing lessons, they could obtain the required resources by searching for keywords. The steward also recommended resources from excellent teachers, reducing the lesson preparation time from 3-4 hours to 1-2 hours.

During teaching, these teachers used tablets to call 3D animations from the steward to demonstrate geometric figures, enhancing students' participation. After class, they pushed personalized materials based on students' situations. After a period of time, the average mathematics score of the class increased by 8 points, the excellent rate rose from 20% to 30%, and the pass rate increased from 60% to 75%. These results highlight the effectiveness of the steward in improving teaching efficiency and quality.

4. Teaching Historical Data Mining: Methods and Value

4.1 Scope and Characteristics of Teaching Historical Data

The scope of teaching historical data is broad, covering teaching resource usage information (such as lesson plan versions and usage frequency), teaching behavior data (such as explanation time and the effect of interactive methods), and students' learning feedback data (such as homework error distribution and classroom performance). Examples include the versions and modification records of A Dream of Red Mansions teaching courseware used by a Chinese teacher in different semesters, as well as students' feedback, and the duration of group discussions and students' participation in a mathematics class.

Teaching historical data has the characteristics of multi-source, dynamic, and complexity. The multi-source feature is reflected in the fact that data comes from multiple channels such as teaching management systems and personal storage platforms, with different formats and structures. The dynamic feature is manifested in the continuous generation of new data from teaching activities, leading to constant data updates. The complexity arises because the data includes both structured data (such as grade tables) and unstructured data (such as teaching reflections), and there are complex correlations between different types of data, which increases the difficulty of data integration and analysis [5].

4.2 Technologies and Methods of Data Mining

Association rule mining aims to discover interesting association relationships between item sets in a dataset, which can identify potential connections between different teaching elements [6]. For example, analysis shows that when teachers adopt the group discussion teaching method and use multimedia resources, students generally achieve higher test scores in related knowledge points. Common algorithms include the Apriori algorithm and the FP-Growth algorithm. The Apriori algorithm is suitable for large-scale data but requires multiple scans of the database, resulting in low efficiency. The FP-Growth algorithm introduces a frequent pattern tree and only needs to scan the database twice, making it more efficient and suitable for large-scale datasets.

Clustering analysis is an unsupervised learning method that groups similar data objects into the same cluster. It can classify teachers' teaching styles and students' learning patterns [7]. By clustering students' learning behavior data (such as learning time distribution and homework habits), students can be divided into types such as hardworking and procrastinating. Common algorithms include the K-means algorithm, hierarchical clustering algorithm, and DBSCAN algorithm. The K-means algorithm is widely used but is greatly affected by the initial cluster centers. The hierarchical clustering algorithm can display the hierarchical structure of data but has a large computational load. The DBSCAN algorithm can handle clusters of arbitrary shapes and is sensitive to data density.

Classification algorithms establish rules through training data to predict the category of new data, which can predict students' academic performance and teachers' teaching effects. For example, using the decision tree algorithm, students' final exam grade levels can be predicted based on their learning foundation and attitude [8]. Common algorithms include decision trees, support vector machines, and naive Bayes. The decision tree algorithm is easy to understand but prone to overfitting. Support vector machines perform well in small-sample and non-linear classification problems but have high computational complexity [9]. The naive Bayes algorithm has high computational efficiency but is highly dependent on data.

4.3 Value of Data Mining for Teaching Resource Supply

Data mining can accurately grasp teachers' teaching needs [10]. Due to differences in subject backgrounds and teaching styles, different teachers have varying resource needs. By analyzing teachers' teaching behaviors, resource usage preferences, and students' feedback in the teaching historical data, their specific needs in different teaching scenarios can be understood. For example, physics teachers who focus on practice need more resources such as experimental videos and simulation experiment software when explaining mechanics knowledge. After data mining identifies this need, relevant resources can be pushed in a targeted manner, improving the pertinence of teaching.

Data mining also helps optimize the allocation of teaching resources. Educational resources are limited; by analyzing the usage frequency and effect of resources, the value of resources can be judged. Efficient resources are promoted, while inefficient resources are optimized or eliminated. For example, if a mathematics teaching courseware is frequently used by many teachers and receives positive feedback from students, it can be further promoted and improved. For materials that are rarely used, their quality is re-evaluated, and improvements or clean-ups are made to avoid resource waste.

Based on the results of data mining, personalized teaching resource recommendation can be realized. Combining teachers' teaching historical data and students' learning data, resources that are relevant to the current teaching content, in line with teachers' teaching styles, and meet students' needs are recommended to teachers. When teachers prepare for teaching a certain chapter, the system can recommend corresponding lesson plans, courseware, and exercises, improving teachers' lesson preparation efficiency, meeting students' personalized learning needs, and promoting the improvement of teaching quality.

5. Resource Matching Mechanism Based on Teaching Historical Data Mining

5.1 Key Elements of Resource Matching

Teachers' teaching styles are important elements of resource matching, as teachers with different styles have different resource preferences. Lecture-based teachers tend to prefer well-organized and logically rigorous materials, such as detailed lesson plans and systematic knowledge explanation documents. Inquiry-based teachers need inspiring resources, such as open-ended problem scenarios and inquiry-based experimental cases. Teachers who are good at multimedia teaching have a high demand for high-quality audio and video resources, while those who focus on classroom interaction need more resources that promote interaction, such as group discussion topics and classroom game materials.

Students' learning needs are the core element of resource matching, which are influenced by their learning abilities, interests, and foundations. Students with strong learning abilities need extended and challenging resources, such as cutting-edge subject materials and high-difficulty exercises. Students with weak learning abilities need basic knowledge explanations and targeted tutoring exercises. In Chinese learning, students interested in literature hope to obtain resources such as classic book interpretations and model essay appreciations. In physics and chemistry learning, students interested in scientific experiments pay more attention to experimental videos and operation guides. In addition, students with different knowledge foundations have different resource needs: those with weak foundations need to start with basic knowledge, while those with solid foundations can access advanced resources.

The type and quality of resources are crucial. There are various types of teaching resources, including document-type (lesson plans, academic papers, etc.), multimedia-type (images, videos, etc.), and software-type (teaching auxiliary software, etc.). Different teaching scenarios and goals require different types of resources. When explaining abstract scientific concepts, animation and video resources can intuitively present principles. When conducting academic research, academic papers and professional books are essential. In terms of quality, resources must be accurate, complete, timely, and applicable. Low-quality resources will interfere with teaching and reduce teaching quality.

5.2 Matching Algorithms and Model Construction

The collaborative filtering algorithm is based on user behavior data. By finding user groups with similar interests and behavior patterns, it recommends resources to target users. In the teaching scenario, by analyzing teachers' teaching historical data (such as resource usage records and evaluations), teachers with similar teaching styles and resource preferences to the target teacher are identified. Then, based on the resource choices of these similar teachers, resources are recommended to the target teacher. This algorithm is divided into user-based and item-based collaborative filtering. User-based collaborative filtering focuses on user similarity, while item-based collaborative filtering focuses on resource similarity. In practical applications, the two are often combined to improve the accuracy and diversity of recommendations.

The content-filtering algorithm matches teachers' teaching needs based on the content characteristics of teaching resources (such as themes, knowledge points, and keywords). It analyzes the text content of resources, extracts key information, and compares it with the demand keywords or knowledge points input by teachers to screen out matching resources. For example, if a teacher needs resources on the knowledge point of "functions", the algorithm will search for resources containing content related to "functions" and recommend them in order of relevance. The advantage of this algorithm is that it is not affected by other users' behaviors and has a good recommendation effect for new teachers or new resources. However, it may ignore characteristics such as resource applicability and teaching effects.

The construction of a resource matching model needs to integrate multiple factors and algorithms. Firstly, the teaching historical data is preprocessed through operations such as cleaning, denoising, and normalization to ensure data quality and availability. Then, key features such as teachers' teaching styles, students' learning needs, and resource attributes are extracted. Next, the collaborative filtering and content-filtering algorithms are combined, and different weights are set to integrate their recommendation results. Machine learning algorithms such as classification and regression can also be introduced to predict teachers' resource needs in different teaching scenarios, optimizing the model and improving the accuracy and intelligence of resource matching.

5.3 Operation and Optimization of the Mechanism

The operation process of the resource matching mechanism includes data collection, demand analysis, resource screening and recommendation, and feedback collection. In the data collection stage, teachers' teaching historical data (such as resource usage records and students' feedback) is collected through the private data steward, serving as the basis for matching. In the demand analysis stage, data mining and AI algorithms are used to analyze the data, explore teachers' teaching styles and students' learning needs, and accurately grasp teachers' resource needs. In the resource screening and recommendation stage, based on the results of demand analysis, matching resources are screened from the resource library using matching algorithms and recommended to teachers in order of matching degree. In the feedback collection stage, teachers' evaluations and feedback on the recommended resources are collected, and the differences between the recommended resources and teachers' actual needs are analyzed to provide a basis for mechanism optimization.

To optimize the mechanism, algorithm parameters and weights need to be adjusted based on feedback data, and the calculation logic and model structure of the algorithms should be improved. For example, if it is found that the recommendation results of the collaborative filtering algorithm are single, the similarity calculation method can be adjusted and more influencing factors can be added. At the same time, the resource library should be optimized: according to teachers' feedback and changes in teaching needs, resources are updated and expanded, low-quality and low-usage resources are deleted, and resources are reclassified and labeled to improve the efficiency of retrieval and matching, ensuring that the mechanism adapts to the needs of teachers and teaching scenarios.

6. Conclusions and Prospects

Theoretically, this study analyzes the theoretical foundation of accurate supply, clarifies the key elements of reform, and compares the current situation at home and abroad. In terms of technical mechanisms, it elaborates on the functions and applications of the private data steward and studies data mining technologies and resource matching mechanisms. In practice, case studies verify the effectiveness of the system in improving teaching efficiency and quality, providing theoretical and practical support for the reform of accurate resource supply.

In terms of technological innovation, the application of deep learning algorithms in data mining will be explored, and emerging security technologies will be studied to ensure data security. In terms of application expansion, the application of the mechanism will be extended to more subjects and educational stages, and it will be combined with emerging teaching models. In terms of theoretical improvement, data mining will be integrated with educational theories, and the psychology of teachers' and students' usage behaviors will be studied to improve the theoretical system.

It is expected that through research and practice, the accurate supply of resources will be promoted to help teachers improve teaching quality and promote the all-round development of students. Information technology will be used to break resource barriers and achieve educational equity. The supply system will be improved to promote the innovation of teaching models, cultivate innovative talents, and serve social development.

Acknowledgments

This work is supported by National College Students' Innovation and Entrepreneurship

Project of Yulin Normal University - Private Data Steward: A Private Capacity Resource Platform Based on Historical Data Mining and 2024 Undergraduate Teaching Reform Project of Yulin Normal University — Research and Practice on the Reform of Network Information Security Talent Cultivation through Industry-Education Integration and Competition-Promoted Learning (2024xjjg03).

References

[1] Gao Yuanyuan, Fu Rong, Wang Dandan, et al. Thoughts on the Teaching of Data Analysis and Mining Course in Medical Colleges and Universities. Computer Education, 2025, (10): 258-263. DOI: 10.16512/j.cnki.jsjjy.2025.10.021.

[2] Xu Tianyi, Yu Ruiguo, Liu Zhiqiang, et al. Interdisciplinary Teaching Reform of Data Mining Course for "Artificial Intelligence +" Talent Cultivation. Computer Education, 2025, (10): 101-107. DOI: 10.16512/j.cnki.jsjjy.2025.10.048.

[3] Liao Jianjun, Zhang Lanfeng, Liao Jiangyan. Problems and Countermeasures in the Teaching of Data Mining Course under the Concept of "Student-Centered". Neijiang Science & Technology, 2025, 46(09): 35-36+115.

[4] Yu Bo, Peng Chen, Liu Cui. Teaching Reform of Data Mining Course under Multi-Modal Fusion. Computer Education, 2024, (04): 38-43. DOI: 10.16512/j.cnki.jsjjy.2024.04.016.

[5] Wang Xinyan, Zhang Xiumei, Chen Cuiping. Analysis of Learning Behavior Characteristics in Smart Classrooms Based on Educational Data Mining Technology. Educational Measurement and Evaluation, 2024, (06): 31-46. DOI: 10.16518/j.cnki.emae.2024.06.003.

[6] Su Xi. Design of Assistant Decision Support Model for Student Evaluation of Teaching Based on Association Data Mining. Wireless Internet Technology, 2025, 22(18): 73-76.

[7] He Ling, Wu Lingda, Cai Yichao. A Survey of Clustering Algorithms in Data Mining. Application Research of Computers, 2007, (01): 10-13.

[8] Liu Hongyan, Chen Jian, Chen Guoqing. A Survey of Data Classification Algorithms in Data Mining. Journal of Tsinghua University (Science and Technology), 2002, (06): 727-730. DOI: 10.16511/j.cnki.qhdxxb.2002.06.005.

[9] Naeem S, Ali A, Anam S, et al. An unsupervised machine learning algorithms: Comprehensive review. International Journal of Computing and Digital Systems, 2023.

[10] Romero C, Ventura S. Data mining in education. Wiley Interdisciplinary Reviews: Data mining and knowledge discovery, 2013, 3(1): 12-27.